SIMPLE MINDREADING ABILITIES PREDICT COMPLEX THEORY OF MIND: DEVELOPMENTAL DELAY IN AUTISM SPECTRUM DISORDERS

Maria Chiara Pino[1], Monica Mazza[1], Melania Mariano[1], Sara Peretti[1], Dagmara Dimitriou[2], Francesco Masedu[1], Marco Valenti[1], Fabia Franco[3]

1. Department of Applied Clinical Sciences and Biotechnology, University of L'Aquila (Italy)
2. Department of Psychology and Human Development, UCL-Institute of Education, London UK
3. Department of Psychology, School of Science and Technology, Middlesex University, London UK

Accepted for publication in *Journal of Autism and Developmental Disorders* on 6th May 2017

DOI 10.1007/s10803-017-3194-1

Available at:

Online First: http://link.springer.com/article/10.1007/s10803-017-3194-1

http://rdcu.be/tlY4

ABSTRACT

Theory of Mind (ToM) is impaired in individuals with Autism Spectrum Disorders (ASD). The aims of this study were to: i) examine the developmental trajectories of ToM abilities in two different mentalizing tasks in children with ASD compared to TD children; and ii) to assess if a ToM simple test known as Eyes-test could predict performance on the more advanced ToM task, i.e. Comic Strip test. Based on a sample of 37 children with ASD and 55 TD children, our results revealed slower development at varying rates in all ToM measures in children with ASD, with delayed onset compared to TD children. These results could stimulate new treatments for social abilities, which would lessen the social deficit in ASD.

Keywords: Autism Spectrum Disorders (ASD); developmental trajectories approach, eyes task, Comic Strip Test (CST); Theory of Mind (ToM)

INTRODUCTION

Social cognition (SC henceforth) includes the cognitive processes based on the ability to process the social world (Mazza et al. 2010). This ability is essential in everyday life as it allows us to predict other people's behaviour and to adequately modify our behaviour accordingly (Vetter et al. 2013). A main component of SC is known as Theory of Mind (ToM), i.e. the ability to understand the mental and emotional states of other people (Mazza et al. 2010). ToM affects the development of social behaviour from birth. The development of ToM ability follows a certain order of sequence in typically developing (TD) children (Peterson et al. 2005; Wellman and Liu 2004). According to Broekhof and collaborators (2015) children with Autism Spectrum Disorders (ASD) displayed the same progressive order but this appeared to be delayed in age of attainment. Thus, individuals with ASD show a developmental pattern in their ToM abilities, but this pattern may not necessarily resemble that of typically developing individuals or be easily compared to other clinical groups such as individuals with mental retardation (Yirmiya et al. 1996).

From birth newborns appear to be predisposed to develop several social skills, such as the capacity to detect social agents, mutual affiliation, and a preference for social patterns compared to non-social patterns (Happè and Frith, 2014). At around 3 months, there is evidence that typical infants develop social reciprocity between self (i.e. the child's own body) and others (such as his/her mother), together with increasing participation in interaction. Happé and Frith (2014) suggest that, by the age of 3 to 6 months, early capacities of emotion processing and sensitivity to social signals emerge, chiefly through gaze processing. A few studies on children later diagnosed with ASD suggest a reduced attention to social scenes at 6 months (Chawarska et al. 2013; Ozonoff et al. 2010). By the age of 2 years, TD children clearly show awareness of the difference between thoughts of the mind and things in the world. At this age, they also understand that a person will feel happy if he/she get what they want, and sad if he/she does not (Astington and Jenkins, 1999). At the same age (about 2 years) children with ASD show attenuated or absent social abilities including pretend play (they prefer the repetitive manipulation of objects), joint attention and social communication. All these socio-cognitive abilities seem to be linked to mentalizing abilities. Prior research on the development of ToM showed that TD children begin to consolidate this ability at the age of 4-5 years (Happè and Frith 2014; Wellman et al. 2001). Around the age of 6 years, this ability starts to become more elaborate, with the development of higher order ToM abilities (Perner and Wimmer, 1985; Vetter et al. 2013), which are necessary to understand complex mental states (Happé 1994). For instance, prosocial behaviour, moderated by the perceived morality of an adult-to-be helped, is strongly evident at this age (Happè and Frith 2014).

In the previous studies ToM abilities was simply considered the ability to attribute mental states to oneself and others, including knowledge, beliefs and intentions (Premack and Woodruff 1978). Recently, the definition of ToM has

expanded to include the understanding of feelings (Mazza et al. 2014; Sebastian et al. 2012; Shamay-Tsoory et al. 2005). Thus, ToM may be described as including two dimensions, respectively, cognitive and affective sub-processes (Amodio and Frith 2006; Mazza et al. 2014; Sebastian et al. 2012; Shamay-Tsoory et al. 2010). Cognitive ToM refers to the ability to make inferences about beliefs, intentions, motivations and thinking, whereas affective ToM is the ability to understand what people feel in specific emotional contest such as their own emotional states are (Sebastian et al. 2012).

In the literature, at least part of the interest in ToM abilities is associated with the idea that individuals with ASD lack ToM abilities and suffer from "mind-blindness" (Baron-Cohen et al. 2015). At a more general level, ASD is characterised by difficulties in two domains: (1) social communication and interaction, (2) repetitive pattern of behaviour (APA 2013; Lai et al. 2014; Mazza et al. 2014). The former symptom could be explained with the impairment in ToM development of individuals with ASD (Broekhof et al. 2015).

Despite reported improvements in ASD symptomatology with age in domains such as motor or language abilities, social abilities remain very limited (Lai et al. 2014). Indeed, children with ASD fail ToM tasks even in middle childhood and adolescence (Peterson et al. 2005). Baron-Cohen and collaborators (1985) found that 80% of children and adolescents with ASD failed a standard false belief test that was passed by TD pre-schoolers as well as by a control group with Down syndrome. Thus, the specific ToM deficit in the individuals with ASD is independent of any general developmental delay (Baron-Cohen 1989), and for this reason Baron-Cohen refers to the ToM deficit in ASD as a case of specific developmental delay. Such deficits are often thought to hinder social inclusion and lead to social isolation in people with ASD (Mazza et al. 2014).

The main purpose of this study was to identify a simple test that could be used for ToM screening in children with ASD of a wide range of cognitive and behavioural functioning. A simplified version of the Eyes-Test (Franco et al. 2014) was used together with Comic Strip test (CST), a well-known test assessing different components of ToM (Sivaratnam et al. 2012). The CST is associated with good receptive language ability about the social stories which the test in based on, whereas the Eyes-Test (Franco et al. 2014) is a simple and fast measure which can to be used with non-verbal or minimally verbal children. Therefore, this test could assist with the early detection of ToM impairments in children with ASD: if failure on a verbal ToM task may be due to the failure to understand mental and emotional states, it may equally reflect children's difficulty to comprehend test stories or instructions (Astington 2001; Frye et al. 1995). There is robust evidence highlighting a strong relationship between language development and ToM (Astington and Jenkins 1999; Colle et al. 2007; Milligan et al. 2007; Sivaratnam et al. 2012, amongst others), for instance showing that children with ASD show impairments in receptive and expressive language (Ambridge et al. 2015). Deficits in nonverbal communicative behaviours impede a fluid development of social interaction and difficulties in nonverbal communication and in social interaction are one of the diagnostic criteria for ASD.

Indeed, most measures used to assess ToM abilities require well-developed expressive and receptive language skills, but many children with ASD have very poor verbal abilities (Colle et al. 2007). Thus, if proven to be a good predictor of more advanced ToM in the typically developing and clinical groups, a single simplified test such as the Eyes-Test (Franco et al. 2014) could substitute complex ToM tools without language mediation, hence offering a tool suitable for work with younger and less able children and based on solid developmental dimensions. For this purpose, in the present study children with ASD were compared to TD children using a developmental trajectories (DT) approach, aiming to construct a function linking performance on specific experimental tasks with chronological and verbal mental age, and then assessing whether such function differs between the TD and disorder groups (Thomas et al. 2009). This approach allows researchers to study separately developmental effects determined by experience (Chronological Age) from effects associated with language functioning per se (Verbal Mental Age). Divergent findings regarding the correlation between ToM abilities and VMA were found. Indeed, some authors found a correlation between these two variables (Charman and Baron-Cohen 1992; Eisenmajer and Prior 1991; Happè, 1994a/b, 1995; Leekam and Pemer 1991; Ozonoff and McEvoy 1994; Prior et al. 1990; Sparrevohn and Howie, 1995); whereas others researchers did not find any correlation (Charman and Baron-Cohen 1992; Leekam and Perner, 1991). These inconsistent results may stem from the fact that many researchers have used tests of single word receptive language skills, such as the British Picture Vocabulary Scale (Dunn et al. 1982) or Peabody Picture Vocabulary Test (Dunn 1981), as a measure of verbal ability. This measure may be too narrow to capture the global verbal ability of different clinical groups (Yirmiya et al. 1996). For this reason, in the present study we have elected to use a complete and complex verbal measure of Mental Age, assessed with the Test for Reception of Grammar - Version 2 (TROG-2; Bishop 2003).

METHOD

Participants

Demographic and clinical information of all the participants are summarized in Table1. Ninety four children participated in this study: 37 individuals with ASD selected by the Reference Regional Centre for Autism of <anonymised> (age range 61-157 months) and 57 TD (age range 60-147 months) recruited from local schools in the same region and matching the ASD group on chronological and verbal mental age (respectively CA and VMA henceforth). The groups presented uneven distributions by gender (respectively, ASD 34 males and 3 females; TD 37 males and 20 females) but mean CA and VMA were not significantly different in the two groups – see Table 1.

The ASD diagnosis was provided by experienced clinicians according to the new criteria of the DSM-5 (APA 2013). ASD diagnosis of patients was confirmed using the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al. 2012). Verbal Mental Age (MA) was assessed with the TROG-2 (Bishop 2003).

- Please insert Table 1 about here –

Procedure

All children were tested individually in a quiet room according to the principles established by the Declaration of Helsinki. Ethical approval was obtained from the hospital's Ethics Committee <anonymised>. The Ethics Committee approved the experimental protocol prior to the recruitment of participants. Both parental informed consent and the child's assent were obtained before participation. Children with autism were tested at < anonymised>, whereas TD children were tested in their nurseries or schools. All the children had Italian as their first language.

There were 12 children with ASD recruited from nurseries, 21 children from primary school settings (8 children the first class; 6 children the second class; 3 children the third class; 4 children the fourth class) and 4 children attended the first class of secondary school. Twenty-five children had assistance from the special needs teachers.. All children with ASD had IQ scores within the normal range although their verbal intelligence quotient index (VIQ) was lower compared to performance intelligence quotient index (PIQ; see Table 1). In addition, results from parental scores on the Vineland Adaptive Behavior Scales, Second Edition (Sparrow et al. 2005) indicated the mean age on verbal abilities of 25.75 months (SD: 4.51) in contrast to motor abilities such as walking of 16.26 months (SD:1.89). All parents reported motor clumsiness and social difficulties.

The order of tests and the experimenter were always the same for all children. A short break was allowed between the completion of the TROG and the start of ToM measures. At the end of the testing session all participants received rewards (i.e., coloured stickers). The first test administered was the TROG-2 (Bishop 2003). We selected TROG-2 because this test assesses the complex and multifaceted nature of receptive language. Indeed, TROG-2 is a standardized measure of receptive syntax, examining 20 specific syntactic constructions in its entirety. Each construction is tested with a block of four items and testing a specific grammatical construction. The participant's task is to select the one drawing out of four choices that corresponds to a sentence read by the examiner. Each block is characterized by a lexical element (noun, verb, adjective) or a grammatical element (word order, function word, inflection), yielding the possibility to evaluate both lexical or syntactic errors (Kover et al. 2014). The total number of blocks passed was used to calculate standard and age-equivalent scores. Many standardized assessments of receptive language, including those that purport to examine specific domains of language (e.g., receptive syntax), are seldom sufficient for characterizing profiles of receptive language impairments because they fail to fully probe mastery of specific linguistic elements or constructions (Oakes et al. 2013; Laws and Bishop 2003).

Two types of ToM measures were collected from the children:

i. *Eyes-Task-simplified* (Franco et al. 2014). The task contained black and white photos of children's eye region, portraying either mental states or primary emotions. The expressions included as primary emotions were happy and surprised (positive/neutral valence), and sad and angry (negative valence), while excited and thinking (positive/neutral valence), and worried and shy (negative valence) were selected to represent mental states. Although in most cases it is unknown what the child was actually feeling at the time of taking the photographs, they were all derived from naturalistic settings (e.g., taken by parents) rather than being posed specifically for an experiment. 56 slides were prepared with each of the eight above-described target emotion/mental states paired with each of the seven remaining emotion/mental states in the series. Testing lasted approximately 15 min, based on randomised presentation of the items. Accuracy was measured on a 0-7 scale, based on the participant's selection of one of two choices in seven comparisons per target emotion/mental state (for further details, see Franco et al. 2014).

ii. *Comic Strip Task (CST)* (Cornish et al. 2010; Sivaratnam et al. 2012). The CST is a recent 21-item measure developed to assess three aspects of ToM. Besides an overall score, it comprises three subscales, testing the understanding of Beliefs, Intentions and Emotions, respectively. Each subscale includes 5 items, each comprising a 5-picture comic strip illustrating everyday social scenarios involving interpersonal interactions familiar to young children. In addition, it contains a control subscale (two items for each subcomponent of ToM), comprising non-social scenarios, albeit in similar settings to the other comic strips. For each item, children are shown three pictures which tell a social story, after which they are presented with two pictures containing alternative endings to the story, and asked to select the one that they think best completes the story. One option indicates a lack of understanding of others' mental states and is scored 0; the other indicates the presence of such understanding, and is scored 1. Each ToM subscale has a maximum score of 5, with a total test score of 15, with higher scores indicating superior ToM ability (Sivaratnam et al. 2012).

Data analysis

Analysis of variance was used to test participants' scores in socio-demographic, clinical and social cognition test measures. In addition, we used regression models to examine the association of ToM trajectories with developmental measures contrasting the ASD and TD groups. First we considered functions that link performance with CA, and then consider developmental functions linking performance with VMA, assessed with a stringent test as TROG. The use of trajectories in the study of developmental disorders has its origin in growth curve modelling (see, e.g., Chapman et al. 2002; Thelen and Corbetta 1994; van Geert 1991) and in the wider consideration of the shape of change in development

(Elman et al. 1996; Karmiloff-Smith 1998; Thomas et al. 2009). For further details of the methodology used here for estimating developmental trajectories, see Thomas et al. (2009). The Statistical Package for the Social Sciences (SPSS) software (version 22; SPSS Inc, Chicago, IL, USA) was used in all statistical analyses.

RESULTS

1. Comparison of ASD to TD groups' performance on social cognition tasks

As shown in Table 2, compared to TD children, children with ASD showed overall a significantly lower score on all Eyes-Test components. On the Comic Strip test, the two groups showed significant differences in the Belief and Emotion components, but no difference in the Intention component was found (see Table 2).

- Please insert Table 2 about here –

2. Developmental trajectories (DT)

Linear regression was used to evaluate how CA and VMA predicted task performance between groups. According to Thomas et al. (2009), the trajectory approach used here has been specifically developed as a statistical method for studying change over time, along with hierarchical linear modelling. Between-groups comparison of DT allowed us to evaluate whether profiles generated from different groups differ significantly in terms of their gradients (indicating the rate of change and direction of the trajectories) or intercepts (showing the onset of the trajectories). We contrasted comparisons between the two groups for trajectories plotted according to CA and VMA within respectively Eyes-Test and CST measures**.**

The regression results showed that CA has a significant relationship with tests used in this study, i.e. Eyes-Test and CST performance in TD children but not in those with ASD (Table 3).

- Please insert Table 3 about here –

VMA showed a significant relationship with both the Eyes-Test and Comic Strip Test performance in TD children; whereas for ASD group the significant relationship between VMA and mental states (both positive and negative) and between MA and intentions component were found (Table 4). No other significant relationships were found.

- Please insert Table 4 about here-

Overall, in both ASD and TD groups, performance was significantly higher in function of CA than VMA, with the exception of Comic Strip/Intention in TD children – see Figure 1a,b.

- Please insert Figure 1a,b about here -

2a. CA-based trajectories of Eyes-Test performance comparing ASD and TD groups

The DT of the Eyes-test scores based on CA are illustrated in Figure 2a.

Positive Primary Emotions-PPE

The regression model explained a significant proportion of variance ($F_{1,92}$=5.218; p=0.002; η^2=.15). The model revealed no significant effect of either Group ($F_{1,92}$= 1.77; p = 0.19; η^2=.02) or Group × CA interaction ($F_{1,92}$= 0.859; p = 0.36; η^2=.009). On the contrary, the main effect of CA was significant ($F_{1,92}$= 5.105; p = 0.02; η^2=.05), showing a lag of 33.7 months, i.e., 2 years and 8 months in the ASD group compared to TD group, indicating impairments in developing this ability .

Negative Primary Emotions-NPE

The regression model explained a significant proportion of variance ($F_{1,92}$=2.702; p=0.05; η^2=.10). The model revealed very small effects, with Group ($F_{1,92}$= 0.59; p = 0.44; η^2=.0007), CA ($F_{1,92}$= 1.59; p = 0.221; η^2= 0.017) and Group × CA interaction effect all non-significant($F_{1,92}$= 0.932; p = 0.337; η^2=.01). For Negative Primary Emotion, the ASD group showed a significant delay of 27 months in comparison to the TD group.

Positive Mental States-PMS

The regression model explained a significant proportion of variance ($F_{1,92}$=7.58; p=0.0001; η^2=.20). The model revealed a significant main effect of Group ($F_{1,92}$= 6.14; p = 0.01; η^2=.06), indicating that the groups are different, as well as an effect of CA ($F_{1,92}$= 10.95; p = 0.001; η^2= 0.11). No significant effect of Group × CA interaction was found ($F_{1,92}$= 0.002; p = 0.968; η^2=.0). For Positive Mental States, the ASD group showed a lag of 58 months, i.e., 4 years and 8 months compared to TD group, indicating that this ability does not develop.

Negative Mental States-NMS

The regression model explained a significant proportion of variance ($F_{1,92}$=9.92; p=0.001; η^2=.25). The model revealed no significant effect of either Group ($F_{1,92}$= 0.96; p = 0.33; η^2=.01) or Group × CA interaction effect ($F_{1,92}$= 3.32; p =

0.072; η^2= 0.04). On the other hand, the main effect of CA was significant ($F_{1,92}$= 12.03; p = 0.001; η^2=.12), showing that two groups differ for CA. For Negative Mental States, the ASD group showed a delay of 14.6 months, i.e., 1 year and 2 months compared to TD group.

2b. CA-based between-group comparison on Comic strip Test performance

The DT of the Comic Strip Test components based on CA are illustrated in Figure 2b.

Beliefs

The regression model explained a significant proportion of variance ($F_{1,92}$=21.89; p=0.001; η^2=.43) and revealed significant main effects for Group ($F_{1,92}$= 13.79; p = 0.001; η^2=.14) showing that two groups differ and CA ($F_{1,92}$= 3.54; p = 0.06; η^2=.04) showing that two groups have significant differences for CA, but no significant interaction between the two factors ($F_{1,92}$= 2.84; p = 0.096; η^2=.031). For the *Beliefs* dimension, the ASD group showed a significant impairment of 84.83 months, i.e., 7 years compared to the TD group.

Emotions

The regression model explained a significant proportion of variance ($F_{1,93}$=6.75; p=0.0001; η^2=.19) and revealed a significant main effect of Group ($F_{1,93}$= 6.58; p = 0.01; η^2=.07) showing that two groups are different as well as a small Group × CA interaction effect ($F_{1,93}$= 3.54; p = 0.05; η^2= 0.04). On the other hand, no significant main effect of CA was found ($F_{1,93}$= 0.54; p = 0.82; η^2=.001). For the *Emotion* dimension, the ASD group showed a lag of 87.5 months, i.e., 7 years and 2 months compared to the TD group, indicating an impairment on this test .

Intentions

The regression model did not explain a significant proportion of variance ($F_{1,92}$=1.97; p=0.12; η^2=.06) but it revealed significant main effects of Group ($F_{1,92}$= 2.66; p = 0.01; η^2=.003) and Group × CA interaction ($F_{1,92}$= 5.13; p = 0.02; η^2= 0.05). On the other hand, no significant main effect of CA was found ($F_{1,92}$= 0.08; p = 0.77; η^2=.001). For the *Intentions* dimension, the ASD group showed a lag of 52.5 months, i.e., 4 years and 3 months compared to TD group.

- Please insert Figure 2a and 2b -

2b. Between-group comparisons of developmental trajectories based on VMA

VMA-based between group comparisons on Eyes-Test performance

For details about DT based on VMA and Eyes-test see Figure 3a.

PPE

The regression model explained a significant proportion of variance ($F_{1,92}$=3.83; p=0.01; η^2=.12). The model revealed no significant main effect of Group ($F_{1,92}$= 1.12; p = 0.73; η^2=.001) as well as the Group × VMA interaction effect ($F_{1,92}$= 0.19; p = 0.66; η^2=.002) and VMA ($F_{1,92}$= 3.21; p = 0.077; η^2=.03), showing a delay of 29.5 months i.e., 2 years and 5 months in the ASD group compared to TD group.

NPE

The regression model explains a significant proportion of variance ($F_{1,92}$=3.28; p=0.02; η^2=.10). The model revealed no significant main effect of Group ($F_{1,92}$= 0.88; p = 0.35; η^2=.010) as well Group × VMA interaction effect ($F_{1,92}$= 0.151; p = 0.7; η^2=.002). On the other hand, the effect of VMA ($F_{1,92}$= 4.98; p = 0.028; η^2= 0.053) was significant showing that the groups differ for VMA. For the Negative Primary Emotion, the ASD group showed significant delay of 38 months, i.e., 3 years and 2 months in comparison to the TD group.

PMS

The regression model explains a significant proportion of variance ($F_{1,92}$=8.45; p=0.0001; η^2=.22). The model revealed no significant main effect of Group ($F_{1,92}$= 2.14; p = 0.15; η^2=.02) as well as Group × VMA interaction effect ($F_{1,93}$= 0.39; p = 0.53; η^2= 0.004). On the other hand, the effect of VMA ($F_{1,92}$= 13.87; p = 0.0001; η^2= 0.13) was significant. For Positive Mental States Primary Emotion, the ASD group showed a lag of 48 months, i.e., 4 years compared to TD group, thus indicating that this ability does not develop.

NMS

The regression model explains a significant proportion of variance ($F_{1,92}$=12.16; p=0.0001; η^2=.29). The model revealed no significant main effect of Group ($F_{1,92}$= 0.001; p = 0.97; η^2=.0001) as well as the effect of Group × VMA interaction effect ($F_{1,92}$= 0.69; p = 0.41; η^2= 0.008). On the other hand, the main effect of CA was significant ($F_{1,92}$= 21.54; p = 0.001; η^2=.19). For Negative Mental States, the ASD group showing a delay of 12 months, i.e., 1 year compared to TD group.

VMA-based between-group comparisons of Comic strip Test performance

For details about DT based on VMA and Comin Strip Test components and total score see Figure 3b.

Beliefs

The regression model explained a significant proportion of variance ($F_{1,92}$=19.81; p=0.001; η^2=.43). No significant main effects for Group ($F_{1,92}$= 1.82; p = 0.18; η^2=.02); VMA ($F_{1,92}$= 2.74; p = 0.10; η^2=.03) and interaction between the two factors ($F_{1,92}$= 0.77; p = 0.38; η^2=.009) were found. For the Beliefs component, the ASD group showed a lag of 98.08 months i.e., 8 years compared to TD group, thus indicating that this ability does not develop.

Emotions

The regression model explained a significant proportion of this variance ($F_{1,92}$=6.84; p=0.0001; η^2=.19). No significant main effects for Group ($F_{1,92}$= 1.14; p = 0.28; η^2=.01) and the effect of Group × VMA interaction effect ($F_{1,92}$= 0.004; p = 0.94; η^2=.0001). On the other hand, a significant effect for VMA ($F_{1,92}$= 4.28; p = 0.04; η^2=.05) was found. For the Emotions component, the ASD group showed a lag of 99.5 months i.e., 8 years and 2 months compared to TD group, thus indicating that this ability does not develop.

Intentions

The regression model explained a significant proportion of variance ($F_{1,92}$=2.79; p=0.04; η^2=.09). Significant main effect of Group ($F_{1,92}$= 7.58; p = 0.007; η^2=.08) and Group × VMA interaction ($F_{1,92}$= 8.37; p = 0.005; η^2= 0.09) were found. On the other hand, no significant main effect of VMA was found ($F_{1,92}$= 0.42; p = 0.51; η^2=.005). For the Intentions component, the ASD group showed an atypical delay of 53.2 months i.e., 4 years and 4 months compared to TD group.

- Please insert Figure 3a and Figure 3b about here –

3. Relationship between Eye-Test and Comic Strip Test

Finally, we used an ANCOVA (analysis of covariance) to study the effect of a "covariate", i.e. a third variable, on the dependent variable when comparing the ASD and TD groups. In this analysis the dependent variables were each component of the Eyes-Task, while the covariate was represented by the Comic Strip components (Belief, Emotion, Intention). The aim of this crucial analysis was to evaluate if a basic mind-reading task such as the simplified Eyes-Test

(Franco et al., 2014) could predict performance on the more complex ToM Comic Strip Test. CA and VMA were initially entered in the model but did not produce any significant effects.

Overall, the results showed that the introduction of the Belief and Emotion CST components as covariates did not delete the differences between groups on each Eyes-Test component because the interaction between group and beliefs/emotions components were not significant; specifically when the covariate is beliefs component the results were: for PPE (group: $F_{1,92}$=1.44; p=0.23; beliefs: $F_{1,92}$=0.30; p=0.58; Group X beliefs: $F_{1,92}$=0.05; p=0.81); for NPE (group: $F_{1,92}$=0.31; p=0.57; beliefs: $F_{1,92}$=2.55; p=0.11; Group X beliefs: $F_{1,92}$=0.08; p=0.78); for PMS (group $F_{1,92}$=0.37; p=0.54; beliefs: $F_{1,92}$=3.27; p=0.07; Group X beliefs: $F_{1,92}$=0.001; p=0.97); for NMS (group $F_{1,92}$=0.22; p=0.88; beliefs: $F_{1,92}$=0.74; p=0.78; Group X beliefs: $F_{1,92}$=2.04; p=0.15). Whereas when the covariate is emotions component the results were: for PPE (group: $F_{1,92}$=0.07; p=0.78; emotions: $F_{1,92}$=1.27; p=0.26; Group X emotions: $F_{1,92}$=0.49; p=0.48); for NPE (group: $F_{1,92}$=0.09; p=0.76; emotions: $F_{1,92}$=0.10; p=0.74; Group X emotions: $F_{1,92}$=0.1; p=0.91); for PMS (group $F_{1,92}$=0.24; p=0.62; emotions: $F_{1,92}$=1.68; p=0.19; Group X emotions: $F_{1,92}$=0.001; p=0.98); for NMS (group $F_{1,92}$=1.39; p=0.24; emotions: $F_{1,92}$=2.31; p=0.13; Group X emotions: $F_{1,92}$=2.68; p=0.10). On the contrary, the introduction of the Intention CST component modified the Eyes-Test results except for NPE. Specifically for PPE (group: $F_{1,92}$=8.63; p=0.004; intentions: $F_{1,92}$=0.79; p=0.37; Group X intentions: $F_{1,92}$=4.63; p=0.03); for NPE (group: $F_{1,92}$=3.007; p=0.08; intentions: $F_{1,92}$=0.69; p=0.41; Group X intentions: $F_{1,92}$=1.22; p=0.27); for PMS (group $F_{1,92}$=14.64; p=0.001; intentions: $F_{1,92}$=2.25; p=0.13; Group X intentions: $F_{1,92}$=8.52; p=0.004); for NMS (group $F_{1,92}$=12.45; p=0.001; intentions: $F_{1,92}$=0.98; p=0.32; Group X intentions: $F_{1,92}$=7.48; p=0.008). Thus, the simple Eyes-Test predicted performance in two components of CST, elucidating significant group differences, which instead are not associated with the Intention component of CST.

DISCUSSION

The present study examined the developmental trajectories of ToM abilities across two different mentalizing tasks in children with ASD compared to TD children, with the hypothesis that the simpler task (suitable for a broader range of children on the autistic spectrum) would predict performance on the more advanced tasks (relying more on verbal reasoning). ToM ability was measured over an 8-year period (age range 5-13 years).

There is agreement in the literature that by around 4-5 years TD children have consolidated a ToM. With development, around 6 -7 years, ToM abilities became more elaborate in TD children as at this age they pass tests of second-order theory of mind (Happè and Frith 2014). In addition, TD 7-year-olds can distinguish, for example, whether a person is lying, joking or being ironic, which are important abilities for everyday social skills (Happè and Frith 2014). Aiming to contribute novel findings to this literature, in the current study we tested the relationship between earlier and

later ToM abilities and we used the DT approach to evaluate how the performance on ToM tests is related with age, and to assess whether this relationship differs between the TD and the ASD groups. In some studies, CA is found to be the most sensitive predictor of developmental trends (e.g., Franco et al. 2014); however, in the present study, both CA and VMA were used to analyse the DT of ToM abilities. According to Thomas et al. (2009), the comparison between disorder and TD groups can be based on both CA and VMA, with the latter being derived from a relevant standardized test. It is important to highlight that the use of VMA is theory-dependent, i.e. it is based on a theory-driven view on which standardized test adequately measures developmental progression in the domain that the experimental task is thought to evaluate (Thomas et al. 2009; Yule 1978). In this study we used the TROG-2 (Bishop 2003), which identifies difficulties in various aspects of grammatical understanding at sentence level (Nordberg et al. 2015). Nordberg and collaborators (2015) showed the correlation between measures of TROG and second-order ToM test.

The results of the present study showed that ASD and TD groups perform significant different in all Eyes-Test components, i.e. both primary emotions (positive and negative) and mental states (positive and negative). Consistently with Franco et al. (2014), children with ASD showed lower scores in the task requiring them to understand emotional and mental states when observing photos of other children's eyes. The ASD group also displayed lower scores in the beliefs and emotions components of CST compared to the TD group, but no significant difference between groups emerged for the CST intentions component. This is consistent with the suggestion that the ability to understand others' beliefs and emotions is more complex with respect to the ability to understand the intentions of other people (Happè and Frith 2014). The latter ability occurs earlier in development as it is already present, to a certain extent, in TD infants (Happè and Frith 2014). Thus, it is possible that in children with ASD this ability becomes available later, and at age of about 6 or 7 years they gain the ability to understand the intentions of other people.

In the linear regression analyses carried out to evaluate the relationship between CA / VMA and ToM measures within each group (TD and ASD), the results showed that CA has a significant relationship with tests used in this study, i.e. Eyes-Test and CST performance in TD children but not in those with ASD. Thus, when controlling for CA, all abilities to understand the mental and emotional states of other people are compromised in children with ASD. It appears that CA does not modulate the performance outcomes in the ASD group on the tasks. CA provides environmental experience in terms of repeated exposure but this does not appear to be specifically associated to performance. The environment of early years does not always allow ASD children to acquire adequate experience about social interaction and behaviour due of the infant social difficulties which in turn limit progress associated with experience, with the addition of potentially interacting variables such as higher levels of autistic traits in families (Hasegawa et al., 2015).

On the contrary, VMA appears to be a good predictor of ToM abilities (Happè 1994). Indeed, scores on the VMA in the ASD group showed a significant relationship between this variable and mental states (both positive and negative), and intentions component (CST). Hobson (1984) showed that children with ASD succeeded on ToM tasks as could be expected from their MA; thus, it can be argued that ToM tests require a specific cognitive ability, which could correspond to VMA.

The results of the DT analysis conducted on the simplified Eyes-Test showed a developmental difference between the ASD/TD groups of 2;08 years and 2;05 years for positive primary emotions on CA and VMA respectively. Similarly, the difference between the two groups was 2;02 and 3;02 years for negative primary emotions respectively on CA and VMA. The developmental delay grew when considering mental states: for positive mental states, compared to the TD-group the ASD-group showed a delay of 4;08 years for CA and 4;00 years for VMA. When considering negative mental states, the ASD delay with respect to the TD group was smaller: 1;02 months for CA and 1 year for VMA, respectively. We thus suggest that the understanding of negative mental state is also a complex ability for children with TD. Previous studies have not been consistent in this matter, with some showing that children identified positive expressions and situations more readily than negative ones (Denham et al. 1990), but Franco et al. (2014) findings showing an advantage for the recognition of negative states and speculating that this may be due to survival advantages offered by this ability.

The DT analysis for CST showed that the delay in the ASD- compared to the TD-group was larger than in the Eyes-Test, which is consistent with the higher level of difficulty and complexity of this test. Indeed, this test requires good expressive and receptive language ability, that is abilities specifically impaired in ASD (Colle et al. 2007). The DT analysis on CST/belief performance showed that the difference between the two groups was 7 and 8 years, for CA and VMA respectively, whereas the ASD-group's performance was impaired on the CST/emotion component with 7;02 on CA and 8;02 on VMA. However, for the CST/intention component the ASD/TD group difference in the development of this ability was less prominent, with 4;03 for CA and 4;04 for VMA, respectively. Interestingly, this suggests that, similarly to the Eyes-Test, development in CST/intention is delayed in the ASD group, rather than lacking as in the CST/belief and emotion component. As already mentioned, the groups were not statistically different in CST/intentions; however the DT approach allowed us to elucidate that the ASD-group has a delay in the development of the ability to recognize other people's intentions; on the contrary the abilities to recognize beliefs and emotions of others is severely compromised in ASD. For example, in our ASD sample, children aged around 60-85 months appeared to be challenged by reading intentions, whereas children around 90-157 months did show good performance in this component and thus they compensated the delay and obtained the same overall level of performance as the TD sample. Thus, the DT

analysis showed that the ASD group has a delay in earlier ToM abilities and an impairment of the complex ToM abilities.

Finally, and very importantly, the results supported the hypothesis that the simplified Eyes-Test (a simple task requiring minimal verbal abilities) could predict the performance on a more complex ToM task such as CST, thus providing researchers and clinicians with an instrument suitable for using not just with high-functioning ASD groups but with the broader spectrum. The current results showed that the introduction of CST components as covariates did not change the pattern of results in Eyes-Test except for the intention component. On the contrary, the interaction between group and intentions component does not affect the negative primary emotions performance. These results could be due to early development of the ability to recognize the intentions of other people compared to other's beliefs and emotions. Thus if a child learned to recognize others' intentions, s/he is more likely to recognize the emotional and mental states through the gaze observation, particularly if those intentions and emotions are negatively valenced, hence raising the possibility of immediate negative consequences for the perceiver.

One important limitation of this study concerns the evaluation of VMA. The results of this study indicate that the VMA is one of the most important predictors of TOM abilities but we know that the VMA is theory-related as it depends by the standardized test that a clinician decides to use. However, we decided to use the VMA instead of general IQ scores, because the latter are not sensitive to ToM abilities. Indeed, studies on children with learning difficulties based on lower IQ scores showed that these children passed ToM tasks, compared to ASD children who did not, for instance, children with Down syndrome with low IQ scores who can perform close to the ceiling scores on ToM tasks (Baron-Cohen et al. 1985; Baron-Cohen 1989; Happè 1994b). Thus, ToM difficulty appears not to be based on the general IQ level, whereas VMA appears a more promising associated measure.

CONCLUSION

Our study highlights the importance to use age-appropriate ToM tests that are compatible with the characteristics of children in different disorder groups; in the case of the broad autistic spectrum, tests relying on minimal verbal ability appear more suitable yet effective in predicting more complex ToM abilities. In addition, the DT results highlighted that the ASD profile is best described in terms of delay in the development of mentalizing abilities rather than by a total lack of ToM. More specifically, the results showed that the development of the CST/intention component is delayed in children with ASD compared to TD children. Moreover, the delay of basic components of ToM does not allow to ASD children to develop complex ToM abilities such as the recognition of beliefs and emotions of other people. Thus early habilitation programs of these abilities could improve social competences, attenuate the social deficit in children with autism and reduce their isolation.

Specifically, habilitation programmes ought to be developed using ToM concepts and divided in several steps in agreement with results of DT analysis: (1) knowledge of the primary emotions with different valences; (2) knowledge of the mental states with different valences; (3) observe the gaze of people or other social cues to learn to interpret someone's intentions/emotions through the interpretation of these cues; (4) awareness that other people can have intentions different by own intentions; (5) ability to discern others' emotions/beliefs based on the knowledge about social context; (6) ability to use the vocabulary of emotions and expression terms in adequate way with own culture; (7) capacity for emotional self-efficacy (Mazza et al. 2010; Kumschick et al. 2014).

In light of the results obtained in this study, the future perspectives in the developmental approach studies in individuals with ASD should be: (1) evaluate the development of ToM precursors in infants and toddlers (2) identify the severity of symptoms on the development of ToM abilities in children with ASD (3) creation of rehabilitation treatments consistent with a delay hypothesis of ToM abilities, and not to a general deficit of these social competences. The creation of evidence-based habilitation treatments should be the main objective of the clinicians and researchers who work in this field.

REFERENCES

Ambridge, B., Kidd, E., Rowland, C.F., & Theakston, A.L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language,* 42(2), 239-73.

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders: DSM-V, 5the dn. Arlington, *VA: American Psychiatric Publishing*.

Amodio, D.M., & Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Review Neuroscience,* 7(4), 268-77.

Astington, J.W. & Jenkins, J.M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology,* 35(5), 1311-20.

Baron-Cohen, S., Leslie, A.M. & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition,* 21(1), 37-46

Baron-Cohen, S. (1989). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, 30(2),285-97.

Baron-Cohen, S., Bowen, D.C., Holt, R.J., Allison, C., Auyeung, B., Lombardo, M.V., et al. (2015). The "Reading the Mind in the Eyes" Test: Complete Absence of Typical Sex Difference in ~400 Men and Women with Autism. *PLoS One*. 27;10(8):e0136521.

Bishop, D.V.M. (2003). *The Test for Reception of Grammar, version 2* (TROG-2). London: Psychological Corporation.

Broekhof, E., Ketelaar, L., Stockmann, L., van Zijp, A., Bos, M.G., & Rieffe, C. (2015) The Understanding of Intentions, Desires and Beliefs in Young Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders,* 45(7):2035-45.

Colle, L., Baron-Cohen, S., & Hill, J. (2007) Do children with autism have a theory of mind? A non-verbal test of autism vs. specific language impairment. *Journal of Autism and Developmental Disorders,* 37(4), 716-23.

Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry*, 74, 195–203.

Chapman, R.S., Hesketh, L.J., & Kistler, D.J. (2002). Predicting longitudinal change in language production and comprehension in individuals with Down syndrome: hierarchical linear modeling. *Journal of Speech Language and Hearing Research,* 45(5), 902-1.

Charman, T., & Baron-Cohen, S. (1992) J Understanding drawings and beliefs: a further test of the metarepresentation theory of autism: a research note. *Journal of Child Psychology and Psychiatry,* 33(6), 1105-12.

Cornish, K., Rinehart, N., Gray, K., & Howlin, P. (2010). *Comic strip task.* Melbourne: Monash University Developmental Neuroscience and Genetic Disorders Laboratory and Monash University Centre for Developmental Psychiatry and Psychology.

Denham, S.A., McKinley, M., Couchoud, E.A., & Holt, R. (1990). Emotional and behavioral predictors of preschool peer ratings. *Child Development,* 61(4), 1145-52.

Dunn, L.M., Dunn, L.M., Whetton, C., & Pintilie, D. (1982). British Picture Vocabulary Test. Windsor, Berks: NFER-NELASON.

Dunn, L.M. (1981). Peabody picture vocabulary test-revised (PPVT-R). Forms L and M. Circle Pines, Minn:American Guidance Service,

Elman, J.L., Bates, E., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). Rethinking innateness: a connectionist perspective on development. Cambridge, MA: MIT Press.

Eisenmajer, R., & Prior, M. (1991). Cognitive linguistic correlates of "theory of mind" ability in autistic children. *British Journal of Developmental Psychology,* 9, 351-364.

Franco, F., Itakura, S., Pomorska, K., Abramowski, A., Nikaido, K., & Dimitriou, D. (2014). Can children with autism read emotions from the eyes? The eyes test revisited. *Research in Developmental Disabilities,* 35(5), 1015-26.

Happé, F., & Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry,* 55(6), 553-7.

Happè, F.G. (1994a). An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders,* 24(2), 129-54.

Happe, F.G. (1994b). Wechsler IQ profile and theory of mind in autism: A research note. *Journal of Child Psychology and Psychiatry*, 35, 1461-1471.

Happe, F.G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66, 843-855.

Hasegawa C, Kikuchi M, Yoshimura Y, Hiraishi H, Munesue T, Takesaki N, et al. Changes in autistic trait indicators in parents and their children with ASD: A preliminary longitudinal study. *Psychiatry Research* 2015 Aug 30;228(3):956-7. doi: 10.1016/j.psychres.2015.05.048. Hobson, R.P. (1984) Early childhood autism and the question of egocentrism. *Journal of Autism and Developmental Disorders,* 14, 85-104

Karmiloff-Smith, A. (1998) Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences,* 2(10), 389-398.

Kover, S.T., Haebig, E., Oakes, A., McDuffie, A., Hagerman, R.J., & Abbeduto, L. (2014). Sentence comprehension in boys with autism spectrum disorder. *American Journal of Speech Language Pathology,* 23(3):385-94.

Kumschick, I.R., Beck, L., Eid, M., Witte, G., Klann-Delius, G., Heuser, I., et al. (2014). READING and FEELING: the effects of a literature-based intervention designed to increase emotional competence in second and third graders. *Frontiers in Psychology,* 5,1448

Lai, M.C., Lombardo, M.V., & Baron-Cohen, S. (2014). Autism. *Lancet,* 8;383(9920):896-910.

Laws, G., & Bishop, D. (2003). A comparison of language abilities in adolescents with Down syndrome and children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 46(6):1324–1339.

Leekam, S., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition*, 40, 203-218.

Liebal, K., Carpenter, M., & Tomasello, M. (2008). Infants produce communicative acts based on the common ground they have shared with their interlocutors. *International Journal of Psychology*, 43 (3–4), 684.

Lord, C., Rutter, M., DiLavore, P.C., Risi, S., Gotham, K., & Bishop, S.L. (2012). *Autism diagnostic observation schedule, 2nd edition (ADOS-2) manual (Part I): Modules 1–4*. Torrence, CA: Western Psychological Services.

Mazza, M., Lucci, G., Pacitti, F., Pino, M.C., Mariano, M., Casacchia, M., et al. (2010). Could schizophrenic subjects improve their social cognition abilities only with observation and imitation of social situations? *Neuropsychological Rehabilitation,* 20(5), 675-703.

Mazza, M., Pino, M.C., Mariano, M., Tempesta, D., Ferrara, M., De Berardis, D., et al. (2014) Affective and cognitive empathy in adolescents with autism spectrum disorder. *Frontiers in Human Neuroscience,* 7;8:791.

Nordberg, A., Dahlgren Sandberg, A., & Miniscalco, C. (2015). Story retelling and language ability in school-aged children with cerebral palsy and speech impairment. *International Journal of Language & Communication Disorders,* 50(6), 801-13.

Milligan, K., Astington, J.W., & Dack, L.A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development,* 78(2), 622-46.

Oakes, A., Kover, S.T., & Abbeduto, L. (2013). Language comprehension profiles of young adolescents with fragile X syndrome. *American Journal of Speech Language Pathology*, 22(4):615-26

Ozonoff, S., Iosif, A.M., Baguio, F., Cook, I.C., Hill, M.M., Hutman, T., et al. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(3):256-66.e1-2.

Ozonoff, S., & McEvoy, R. E. (1994). A longitudinal study of executive function and theory of mind development in autism. *Development and Psychopathology*, 6, 415-431.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that . . .": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology,* 39, 437-471.

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theoryof- mind development for children with deafness or autism. *Child Development*, 76(2), 502–517.

Premack, D., & Woodruff, G. (1978). Chimpanzee problem-solving: a test for comprehension. *Science,* 202(4367), 532 5.

Prior, M. R., Dahlstrom, B. & Squires, T.L. (1990). Autistic children's knowledge of thinking and feeling states in other people. *Journal of Child Psychology and Psychiatry*, 31, 587-601.

Sebastian, C.L., Fontaine, N.M., Bird, G., Blakemore, S.J., Brito, S.A., McCrory, E.J. et al. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitive and Affective Neuroscience,* 7(1), 53-63.

Shamay-Tsoory, S.G., Tomer, R., Berger, B.D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired "affective theory of mind" is associated with right ventromedial prefrontal damage. *Cognitive and Behavioural Neurology,* 18(1), 55-67.

Shamay-Tsoory, S.G., Harari, H., Aharon-Peretz, J., & Levkovitz., Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex,* 46(5), 668-77.

Sivaratnam, C.S., Cornish, K., Gray, K.M., Howlin, P., & Rinehart, N.J. (2012). Brief report: assessment of the social-emotional profile in children with autism spectrum disorders using a novel comic strip task. *Journal of Autism and Development Disorders,* 42(11), 2505-12.

Sparrevohn, R. & Howie, P. M. (1995). Theory of mind in children with autistic disorder: Evidence of developmental progression and the role of verbal ability. *Journal of Child Psychology and Psychiatry*, 36, 249-263.

Sparrow, S.S., Cicchetti, D.V., & Balla, D.A. (2005). Vineland Adaptive Behavior Scales, Second Edition (Vineland-II). Pearson.

Thelen, E., & Corbetta, D. (1994). Exploration and selection in the early acquisition of skills. *International Review of Neurobiology,* 37, 75–102.

Thomas, M.S.C., Annaz, D., Ansari, D., Serif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research,* 52, 336-358.

van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review,* 98, 3–53.

Vetter, N.C., Leipold, K., Kliegel, M., Phillips, L.H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology,* 19(6), 615-29.

Wellman, H.M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development,* 72(3), 655-84.

Wellman, H.M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75(2), 523–541.

Yirmiya, N., Solomonica-Levi, D., Shulman, C., & Pilowsky, T. (1996). Theory of mind abilities in individuals with autism, Down syndrome, and mental retardation of unknown etiology: the role of age and intelligence. *Journal of Child Psychology and Psychiatry*, 37(8), 1003-14.

Yule, W. (1978). Research methodology: *What are the "correct controls"?* In M. Rutter & E. Schopler (Eds.), Autism: A reappraisal of concepts and treatment (pp. 155–162). New York: Plenum.

Table 1: Demographic data of the sample and clinical information concerning the ASD group

	TD group (n = 57) *Mean (SD)*	ASD group *(n = 37)* *Mean (SD)*	F *DF = 1,93*	*p*
Demographic data				
Chronological Age	85.65 (26.17)	95.91 (24.9)	2.54	0.083
Verbal Mental Age	90.74 (28.86)	83.13 (27.40)	1.60	0.209
Gender	37 M; 20 F	34 M; 3 F	-	-
Education in years	2.47 (1.94)	2.94 (2.13)	1.22	0.271
VIQ	-	82.48(18.24)	-	-
PIQ	-	92.19(17.20)	-	-
TIQ	-	82.11(17.74)	-	-
Clinical information				
ADOS-social communication and social interaction	-	9.48 (2.93)	-	-
ADOS- Repetitive and Stereotyped Behaviours	-	1.29 (1.2)	-	-
ADOS Total scores	-	10.78 (3.11)	-	-

Legend: VIQ= verbal intelligence quotient; PIQ= performance intelligence quotient; TIQ= total intelligence quotient

Table 2: Summary of TD and ASD groups' performance in two social cognition measures.

	TD group *Mean (SD)*	ASD group *Mean (SD)*	F *DF = 1,93*	*P*	Effect Size η^2_p
Eyes-Test					
Positive Primary Emotion-PPE	6.06 (0.76)	5.50 (1.27)	7.02	0.009	0.14
Negative Primary Emotion-NPE	6.33 (0.63)	5.97 (1.06)	4.17	0.044	0.10
Positive Mental States-PMS	5.88 (0.95)	5.11 (1.45)	9.37	0.003	0.19
Negative Mental States-NMS	5.52 (0.97)	4.91 (1.14)	7.48	0.008	0.14
Comic Strip Test					
Beliefs	3.40 (0.66)	3.81 (0.99)	53.39	0.001	0.37
Emotions	4.49 (0.66)	3.81 (0.99)	15.49	0.001	0.20
Intentions	3.72 (1.34)	3.71 (1.36)	0	0.993	0
Total score	11.61 (1.87)	9 (2.38)	34.72	0.001	0.28

Table 3: Summary of linear regression analyses evaluating the relationship between Chronological Age and ToM measures within each group (TD and ASD)

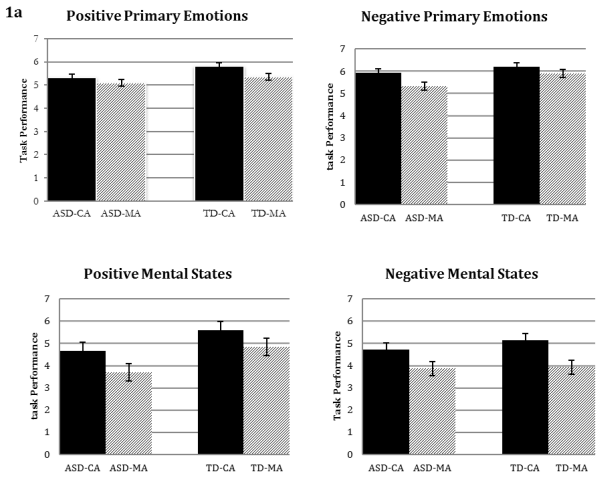
	β	R^2	F	p
TD Group				
Eyes-task				
Primary Emotions				
Positive	0.013	0.197	13.53	0.001
Negative	0.008	0.102	6.25	0.01
Mental States				
Positive	0.015	0.178	11.87	0.001
Negative	0.021	0.330	27.13	0.0001
Comic Strip test				
Beliefs	0.018	0.113	6.75	0.01
Emotions	0.007	0.07	4.96	0.03
Intentions	0.014	0.08	4.51	0.04
ASD Group				
Eyes-task				
Primary Emotions				
Positive	0.05	0.01	0.39	0.53
Negative	0.01	0.001	0.2	0.88
Mental States				
Positive	0.16	0.073	2.76	0.10
Negative	0.007	0.02	0.72	0.36
Comic Strip test				
Beliefs	0.001	0.001	0.025	0.87
Emotions	0.006	0.02	0.76	0.38
Intentions	0.011	0.04	1.47	0.23

Table 4: Summary of linear regression analyses evaluating the relationship between Mental Age and ToM measures within each group (TD and ASD)

	B	R^2	F	P
TD Group				
Eyes-Test				
Primary Emotions				
Positive	0.008	0.102	6.21	0.01
Negative	0.006	0.069	4.08	0.05
Mental States				
Positive	0.013	0.157	10.24	0.002
Negative	0.019	0.337	28.002	0.0001
Comic Strip test				
Beliefs	0.013	0.059	3.35	0.05
Emotions	0.007	0.082	4.73	0.03
Intentions	0.013	0.059	3.31	0.05
ASD Group				
Eyes-task				
Primary Emotions				
Positive	0.005	0.012	0.42	0.51
Negative	0.008	0.04	1.61	0.21
Mental States				
Positive	0.018	0.12	4.76	0.03
Negative	0.013	0.103	4.02	0.05
Comic Strip test				
Beliefs	0.004	0.012	0.42	0.52
Emotions	0.006	0.029	1.04	0.31
Intentions	0.18	0.13	5.11	0.03

Figure 1

Within-group comparison between CA and MA for a each eyes-test component evaluated, and b each CST component evaluated. All comparisons were significant with $p \leq 0.001$.



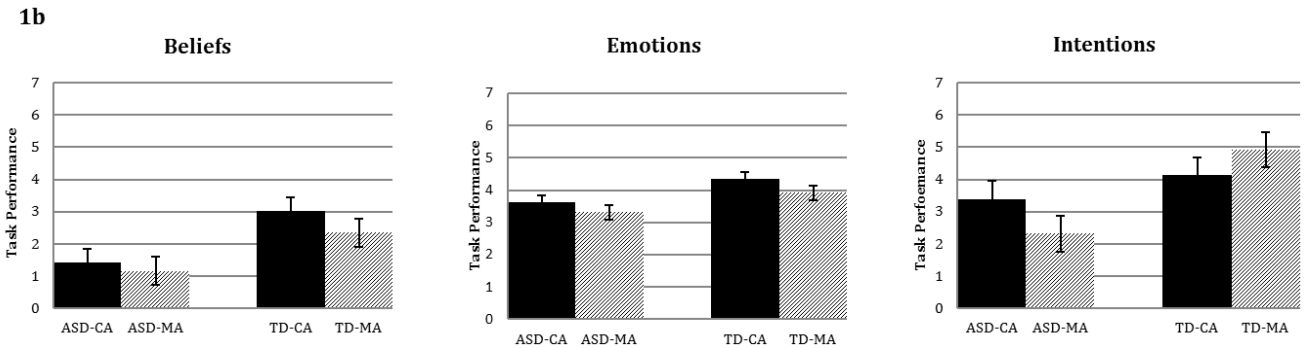


Figure 2

Between-groups comparison of developmental trajectories based on CA: a on the eyes-test and b on the comic strip task.

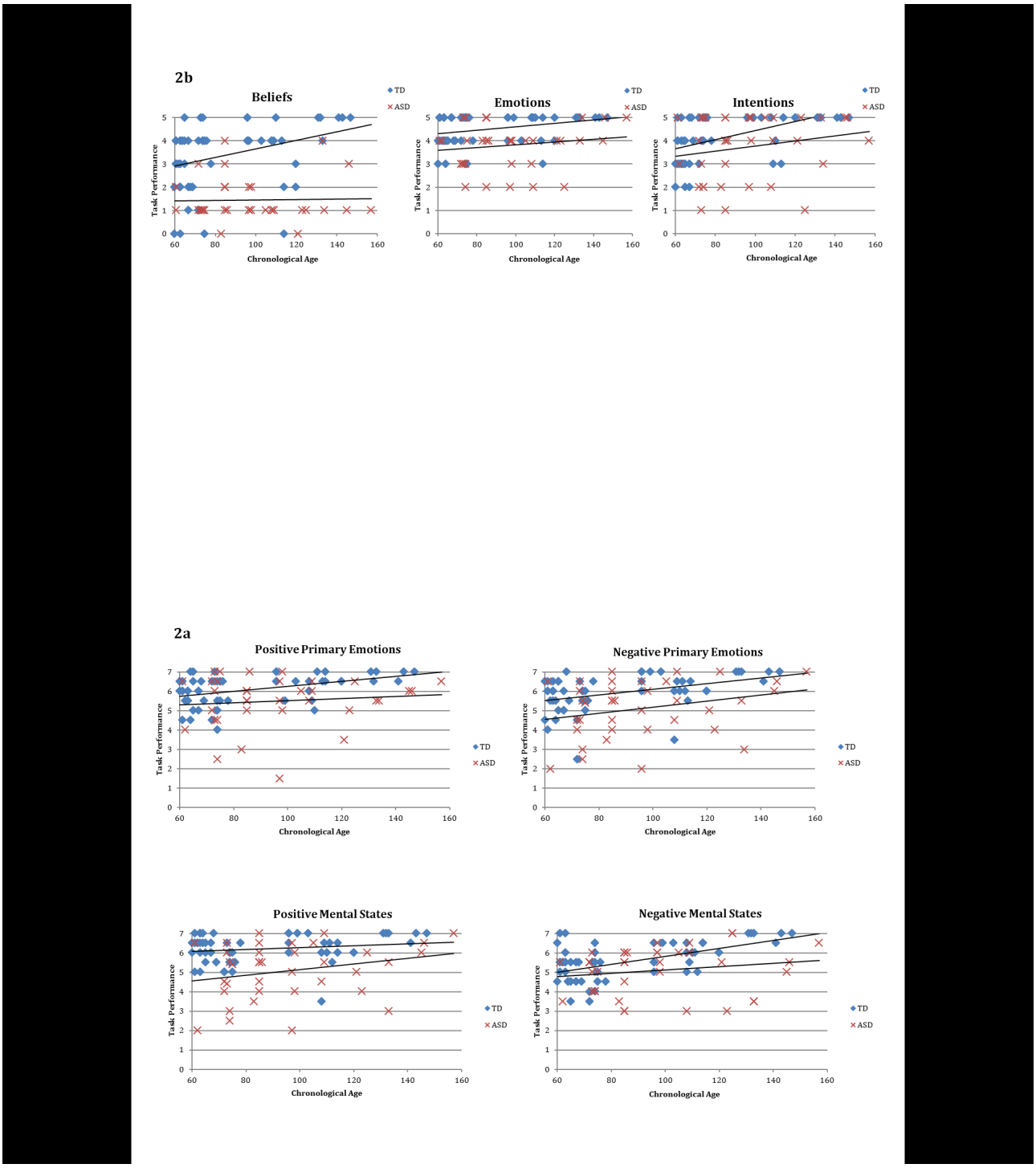


Figure 3

Between-groups comparison of developmental trajectories based on MA: a on the eyes-test and b on the comic strip task.